Description

Method and network node for self-regulating, autonomous and
decentralized traffic distribution in a multipath network

5

The invention relates to a method for redistributing traffic in
response to a traffic overload or a link failure and a node for
a packet-based network with multipath distribution.

10 Possibly the most important development at present in the field
of networks is the development of new technologies to make
packet-oriented data networks and thereby primarily the
internet suitable for new services. These new services relate
primarily to so-called real-time traffic, e.g. voice data and
15 video data.

The success of the internet as the most widely used medium for
sending or transmitting data is largely due to its flexibility.
IP (Internet Protocol) technology is used for so-called
20 destination based routing. IP packets contain an origin address
and a destination address, based on which local routing
decisions are made. This is also referred to as routing on a
per hop basis, in other words routers forward to a subsequent
node (next hop) based on the destination address specified in
25 the data packet. Conventional IP technology has the significant
advantage that states do not have to be maintained in the
network and paths do not have to be determined beforehand for
the transmission of data packets. This has to be offset against
the disadvantage that the packets can only be switched on a
30 "best effort" basis, i.e. there is no guarantee of quality of
service.

The most commonly used routing protocol for an autonomous
network domain or autonomous system at present is the so-called
OSPF (Open Shortest Path First) protocol, which utilizes a
distributed routing approach. The topology of the network is
5   thereby known to the every node and is distributed by so-called
link state advertisement messages to all the nodes in the
network. The shortest path in each instance is calculated
autonomously in every node for all the destinations known to
said node based on link metrics. Link failure results in re-
10   calculation of the paths.

Two main approaches are followed to improve the quality of
service of IP networks in respect of new services with quality
of service requirements. With one approach more than one path
15   is established between a source and a drain or an origin and a
destination, in order to be able to respond to bottleneck
situations or fault situations by distribution via different
paths. The routing protocol OSPF has been extended to this end
by introducing the Equal Cost Multipath (ECMP) concept, which
20   allows the simultaneous use of a number of paths of equal
length, in the sense of a metric, between a source and a drain.
The method makes it possible to distribute the traffic to paths
of the same length. An extension of this concept can be
understood to be the optimization of link usage based on
25   measured or calculated network loads using traffic engineering
mechanisms, i.e. mechanisms to control and regulate the packet
traffic. The so-called OSPF-TE (TE: Traffic Engineering) method
tries to allow regular distribution of the traffic in the
network by adjusting the link metrics distributed in the
30   network as a function of the usage of the respective link. The
multipath routing methods are typically based on a central load
distribution calculation, i.e. a decision is made at a central

point or by a central entity about how the packets should be distributed.

The second approach is the central determination of paths
through the network along with permanent reservation of the required bandwidth. The most important concept, involving the switching of paths, is the so-called Multi Protocol Label Switching (MPLS) protocol or concept. Virtual paths through the network are defined in the context of this concept. Incoming packets are allocated to virtual paths at the input router of an MPLS domain. Appropriate global selection of the paths and the assignment of packets to the paths allow regular distribution in the network.

Such approaches to improving the transmission quality of conventional data networks to some extent at least cancel out the advantages of conventional data networks, namely their flexibility and low level of complexity.

The object of the invention is to specify a method for multipath routing, with which traffic distribution or load distribution is optimized whilst avoiding the disadvantages of conventional methods.

The object is achieved by claims 1 and 9.

The method is based on a packet-based network, in which there is provision for multipath routing or the distribution of packets to different paths. The network is made up of nodes and links (also referred to as edges or connection segments). It is possible, for some of the nodes at least, to send incoming packets to their destination on a number of different paths. The term array of paths or distribution array is used below for

the different alternatives or different options for a node to forward a data packet to a subsequent link as a function of the destination of the data packet. The array of paths for a node in respect of a destination comprises the links out of the

5   node, via which packets can be routed to the destination. The paths to the destination do not necessarily have the same length in the sense of a metric. A destination is generally defined by the routing criterion for the individual nodes specified in a routing table. This is not necessarily identical

10   to the destination address generally specified in the packet but can, for example, comprise part of the destination address. In the case of IP networks the nodes or routers generally use the part of the address describing the network, in which the IP network end point (e.g. host gateway) is located, which

15   corresponds to the full destination address of the packet.

According to the invention, nodes in the network independently or autonomously distribute the packets or the load to the links of an array of paths. This load distribution is also changed

20   autonomously by the nodes. A node monitors the traffic load or packet rate it routes to a link. If an overload occurs in a link, this is detected by the node and an overload protection measure is then initiated. The primary overload protection measure is to redistribute traffic from the overloaded link to

25   other links in the associated array of paths. A response to a link failure by redistributing to links of the associated array of paths that have not failed can be provided for correspondingly.

30   The invention has the advantage that the overload protection response by the node is autonomous, i.e. redistribution takes place locally. There is no need for a complex signaling operation or the propagation of messages through the entire

network of nodes, which would very significantly increase the
complexity of the routing protocol. Nor is there a need for a
central entity, either for control purposes or to determine
weightings if node weightings should have to be recalculated.
5   Instead the weightings for routing purposes along an array are
once again determined by the respective nodes. A central
control entity is a weak point in respect of any fault in said
entity and this is avoided in the present invention.

10 The inventive redistribution or allocation of the traffic to
different paths can be infinite in the sense that any
allocation ratios of packets to links of an array of paths are
possible. This contrasts with various current multipath routing
methods (e.g. ECMP), which specify a fixed allocation ratio.
15 Optimum allocation ratios can for example be determined using
the traffic matrix, i.e. the measured or calculated traffic
loads on the links. Alternatively it is possible to optimize
allocation ratios by means of a control process. The modified
link loads in the event of traffic redistribution are thereby
20 related to the traffic redistribution measures and the
distribution measures to avoid overload situations are thereby
determined.

According to one development, load redistribution is carried
25 out by the individual nodes or by one individual node in
stages. No further traffic redistribution measures are thereby
implemented after the end of a stage during a waiting period or
a time interval. After the end of the time interval any next
stage that may be required for traffic redistribution is
30 started. This allows the nodes following after a node
initiating a protection measure to resolve any overload
resulting at them due to the redistribution and to redistribute

the traffic in turn. Also gradual load adjustment counteracts
system oscillation.

The subject matter of the invention can be developed for
5   situations, in which the overload cannot be resolved by
redistribution within an array or if a link fails the data
packets cannot be distributed to the other links of the
associated array of paths, without generating an overload on
one of these links. To this end an overloaded node sends a
10   message to one node upstream from it at least, prompting said
node to carry out a redistribution operation for its part, such
that fewer data packets are transmitted to the node in
question. Upstream thereby relates to the paths of the packets
distributed on the distribution arrays affected by the overload
15   or link failure. This method can be extended recursively or in
a cascaded fashion to the upstream nodes in front of the node
in question, such that every node that receives a message
requesting redistribution but cannot achieve this by
redistribution within the array or arrays for its part sends
20   messages to nodes upstream from it, so that these for their
part reduce the traffic by redistribution. It is thereby
expedient for a message prompting redistribution to contain
information about the required traffic reduction, e.g. its
nature and scope. This method generally reaches a limit when an
25   edge node in the network receives a message to reduce the
overload. For edge nodes can generally not communicate with
nodes in another network using the protocol used within the
network to transmit messages. To reduce the traffic going away
from it, such an edge node can reduce the traffic it allows
30   into the network as part of a network access control, to bring
about the required traffic load reduction in this fashion.
Access controls are frequently used in data networks, which are

intended to ensure transmission in compliance with quality of
service features.

The subject matter of the invention also includes a network
5   node or router, which is equipped [with means to identify] a
traffic overload and to redistribute traffic to other links
going out from nodes. Such a node can also be provided with
means to send messages, said messages being able to bring about
traffic redistribution in upstream nodes. These means for
10   example include software modules, which support the protocol
used to send the messages.

The subject matter of the invention is described in more detail
below based on an exemplary embodiment with reference to two
15   figures, in which:

Figure 1   shows a network segment with nodes and links

Figure 2   shows an associated table with details of
20   distribution arrays assigned to the nodes shown in the network
segment.

Figure 1 shows nodes 1 to 6 and links (connections between the
nodes) LIJ, I, J out from (1, 2, 3, 4, 5, 6, 7). A direction is
25   thereby assigned to a link. For example the link L31 goes from
the node 3 to the node 1. The overall traffic on the associated
physical link between the nodes 3 and 1 is therefore made up of
the traffic on the link L31 and the link L13, if there is
provision for sending packets from the node 1 to the node 3.
30   For the sake of simplicity links are only shown with one
direction in the figure. Figure 2 shows a table with different
arrays of paths going out from the node 3. The first column D
shows the destination or drain of packets forwarded from the

node 3. For simplicity it is assumed that each of the other
nodes shown is a possible destination for the packets sent. The
second column shows the arrays of paths WF to the respective
destinations. For example an array comprising three links is

5 defined to the destination defined by the node 1, one link
leading directly to the node 1 and the other two respectively
via the nodes 2 and 7. The distribution ratio R or load
distribution on these distribution arrays or arrays of paths is
shown in the third column of the table. 70% of the traffic is

10 thus routed directly to the node 1, while the other 30% is sent
via the node 2 or the node 7 to its destination, the node 1.
The distribution arrays to the nodes 2, 4, 5, 6 and 7 are shown
in a similar fashion. The distribution ratio was determined
according to the path priorities, with direct links having a

15 higher distribution weighting. If the link L31 between the node
3 and the node 1 now fails, according to the invention the node
3 can distribute the traffic to the node 1 via the nodes 2 and
7 by changing the distribution weightings accordingly. As the
majority of the traffic was sent via the link L13 between the

20 node 1 and the node 3, redistribution can cause an overload on
the links L37 between the node 3 and the node 7 and L32 between
the node 3 and the node 2. To reduce this overload, the node 3
sends a message to the node 4 and the node 5, to reduce the
traffic routed to it. On receipt of this message the nodes 4

25 and 5 change their traffic distribution for packets to the
destination defined by the node 1. For example the node 5
routes a larger proportion of the traffic to the node 1 via the
nodes 6 and 7 instead of via the node 3. And the node 4
increases the proportion of traffic routed via the node 2 to

30 the node 1. This results in a reduction of the load on the
links L37 and L32 between the nodes 3 and 7 and 3 and 2, in
that upstream nodes, namely 5 and 4, route more traffic via the
links L54, L42, and L21, as well as L56, L67 and L71. In actual

networks with very many more nodes, the node 5 would generally have a node upstream from it, to which it could in turn send a message, if the measure taken by it would not result in the necessary reduction in traffic volume.

Claims

1.   Method for redistributing traffic in response to a traffic
overload or link failure in a packet-based network made up of
5   nodes (1, 2, 3, 4, 5, 6, 7) and links (L31, L32, L37, L71, L21,
L42, L43, L67, L53, L54, L56) with multipath distribution of
packets, with packets with the same destination being able to
be distributed to at least two links, which are associated with
an array of paths (L31, L32, L37) assigned to the destination,
10 at least for some of the nodes,
according to which
- a node (3) determines when the traffic load on a link (L31)
going away from it exceeds an upper limit or a link (L31) going
away from it fails,
15 - the node (3) redistributes the traffic load independently of
external control entities, by routing at least some of the
packets, which would have been routed via the outward link
(L31) had the load distribution not been changed, to one or
more other links (L32, L37) going away from the node (3), which
20 are assigned to the same array of paths (L31, L32. L37).

2.   Method according to claim 1,
characterized in that
- the traffic load redistribution is carried out in stages in
25 respect of proportions of the overload and
- at the end of a stage there is a waiting period of a time
interval before the next stage is initiated.

3.   Method according to claim 1 or 2,
30 characterized in that
- if the traffic load does not drop below the upper limit as a
result of the traffic redistribution without a different upper
limit being exceeded on a link going away from the node, a

message is sent to a node (4, 5) that is upstream in respect of packets distributed via the arrays of paths (L31, L32, L37), which causes a traffic load redistribution at this upstream node (4, 5) to reduce the traffic load sent by the [lacuna] to

5 the [lacuna] and distributed via the arrays of paths (L31, L32, L37).

4.   Method according to claim 3,
characterized in that

10 - a message is sent to all nodes (4, 5) directly upstream in respect of packets distributed via the arrays of paths (L31, L32, L37) for the purpose of traffic redistribution at these upstream nodes to reduce the traffic load sent by the [lacuna] to the [lacuna] and distributed via the arrays of paths (L31,

15 L32, L37).

5.   Method according to claim 3 or 4,
characterized in that
- the message contains information about the required traffic

20 load reduction.

6.   Method according to one of claims 5,
characterized in that
- a notified upstream node (4, 5), which cannot achieve the

25 required traffic load reduction for the node in question by means of traffic load redistribution without causing a further upper limit for a traffic load to be exceeded, for its part sends a message to a node upstream from it, prompting said node to carry out the required traffic load reduction by traffic

30 redistribution or by sending messages to upstream nodes.

7.   Method according to claim 6,
characterized in that

- upstream nodes, which receive a message from the node in question (3) or from an upstream node (4, 5) either carry out the traffic load reduction required according to the message received by traffic load redistribution or send at least one
5   message to an upstream node to reduce the traffic load,

8.   Method according to one of claims 3 to 6, characterized in that
- an edge node, which receives a message to reduce the traffic
10   load at an upstream node, brings about the required traffic load reduction by reducing the limits for controlling access to the network.

9.   Node (3) for a packet-based network with multipath
15   distribution, with
- means for identifying a traffic overload on a link going away from it and
- means for redistributing traffic to other links going away from the node (3) and associated with the same distribution
20   array.

10.   Node (3) according to claim 9,
with
- means for sending messages to a node (4, 5) that is upstream
25   in respect of packets sent via the distribution arrays (L31, L32, L37) to request a traffic reduction in the traffic transmitted from the upstream node (4, 5).